TECHNICAL BULLETIN 62-18

JUDGEMENTS OF SIMULATED TARGET AXIS ANGLES BY MEANS OF A PRESCRIBED CATEGORY SYSTEM

E. G. AIKEN
W. E. COOLEY

DECEMBER 1962

BUREAU OF NAVAL PERSONNEL

Submitted by

E. I. Jones, Ph.D., Director, Navy Training Research Laboratory
E. E. Dudek, Ph.D., Technical Director
CDR H. B. Boaz, USN, Officer in Charge

U.S. Naval Personnel Research Activity
San Diego 52, California

Approved by

S. Friedman
Head, Psychological Research Branch

CAPT R. J. Baxter, USN
Director, Personnel Research Division

Bureau of Naval Personnel
Washington, D. C.

BRIEF

During the summer of 1961, two groups of students at the Fleet Anti-Submarine Warfare School, San Diego, were tested on their ability to make target axis angle estimates from simulated materials utilizing the angular categories in the NEL classification system. One group received a 30 minute familiarization training period while the other was simply introduced to the task prior to testing. It is concluded that neither training nor the scope quadrant in which the stimulus appears significantly affect accuracy while the angular category does (pp. 9 and 11). The dominant error tendency is toward overestimation of angular size, except at the most acute angular category (p. 10). Accuracy of classification within the Navy Electronics Laboratory, San Diego, system ranges from 50.2% to 89.2% depending on category, with an over-all accuracy of 68.0% (p. 12). The majority of errors (92%) are of only one category in extent (p.12).

CONTENTS

TABLES

FIGURES

JUDGMENTS OF SIMULATED TARGET AXIS ANGLES BY MEANS OF A PRESCRIBED CATEGORY SYSTEM

A. Background

The general question of the ability of human observers to estimate angle size is an old one. A wide variety of techniques have been employed. Some researchers have required that estimates in degrees be made (1, 2, 4, 7, 9, 11). Others have required that a variable be set equal to a standard or have utilized both estimation and adjustment approaches in combination (5, 6, 8, 10). The majority of the studies requiring numerical estimation have employed visual presentation, although at least one writer reports use of kinesthetic input for the estimate (2). Due to the considerable disparity in apparatus, stimulus properties, response requirements, etc. in the literature, it is not possible to draw conclusions, unambiguously, concerning either the amount of error in degrees to be anticipated or its predominant direction.

Typically, the problem of target classification by operators using various information displays requires several fairly basic discriminations or judgments. One of these is an operational requirement for the estimation of angle size by some means or another. Reasonably accurate estimations of acute angles forms an integral part of the Navy Electronics Laboratory (NEL) classification system (3). It is of considerable importance, therefore, to know something about the limits of this sort of judgmental accuracy, as well as the variables which may influence it. In a pilot study involving four subjects, photos from both real and simulated PPI scopes were judged for target axis angle, and median average errors of 13.9 and 2.98 degrees were obtained, respectively (4). It was speculated that this sort of skill might be trainable. The question of the possible effect of the particular scope quadrant in which the axis angle might appear was not investigated. Almost no data exist on the accuracy of the estimation of angle size when a prescribed category system, such as that developed at NEL, has been imposed on the subject in advance. Also, there is evidence that the cursor-target type of stimulus presentation poses a very different problem from the typical procedure of other investigators in which the angle is judged relative to verbally or visually specified coordinates of a circle. In addition, the wide variations in category size in the NEL system pose specific questions that are difficult to answer by generalizing from the existing evidence. For these reasons the study here reported was conducted.[1]

[1]The study here reported is part of a series of investigations relating to the ability of human observers to accurately assess certain types of displayed information. The plan of the series aims at measuring first the fundamental psychophysical functions involved, with close control over stimulus and response features. This stage will define the limits of the ability under idealized

Utilizing simulated target axis angles, the investigation was designed to answer the following questions:

1. How accurately can observers be expected to judge axis angles utilizing the NEL category system?

2. What is the relationship between accuracy in angular judgment and the angle category to which the physical stimulus belongs?

3. What is the relationship between accuracy of angular judgment and the scope quadrant in which the angle is displayed?

4. What effect does training in axis angle judgment have on accuracy?

5. What is the relationship between the constant errors in axis angle judgment, if any, and the angle category, scope quadrant, and training variables?

B. Procedures

1. Subjects

The sample consisted of 60 students in the Basic Surface Sonarman Course C-560G at the Fleet Anti-Submarine Warfare School, San Diego. The subjects were randomly assigned to two experimental conditions.

2. Apparatus and Test Materials

To provide maximum control of the experimental variables and to approach the limit of subject capacity at the task, the test stimuli were simulated by making them up in line drawing form, photographing them on 35 mm film, and then projecting them at the time of the test. By this means, the axis angle of each stimulus could be specified precisely; something quite difficult to do with real PPI photos. The test materials constituted an uncluttered, "noise-free" simulation of the operational task of discerning cursor-target angle. Real PPI displays contain the additional problem of deciding on the major axis of the target pip, which is often quite ambiguous. It can reasonably be assumed, therefore, that performance will be degraded in the operational setting from what is here reported. Figure 1 presents a typical test item.

---

circumstances. This will be followed by investigations in which the operational situation is more and more closely approximated. The study here reported is an intermediate one, involving both controlled simulation and operational features.

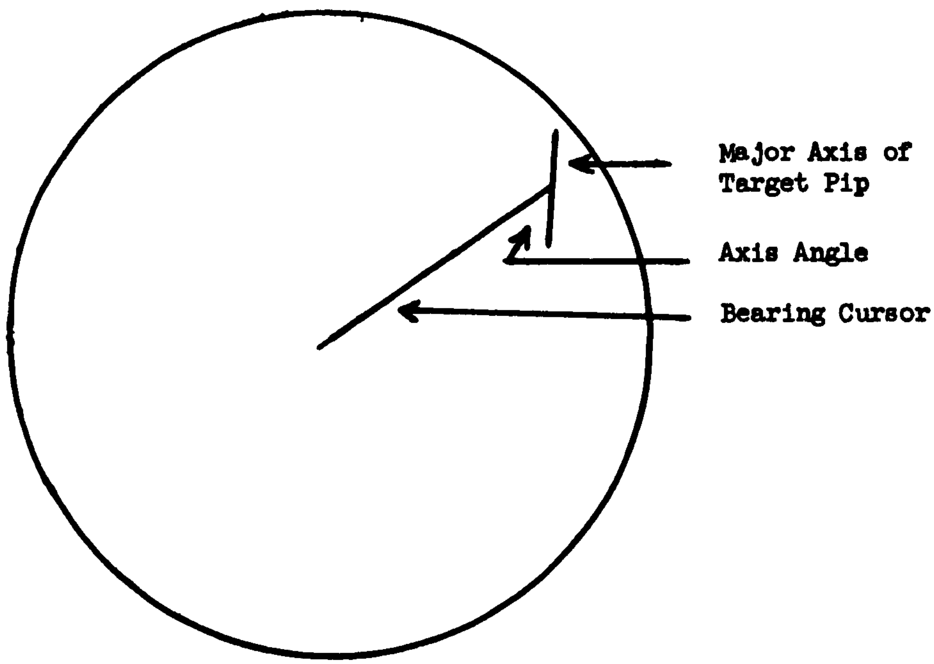

Fig. 1. Representative test item

Axis angle is defined as the extent of the acute angle formed by the bearing cursor and the major axis of the target pip. It follows that it can vary from 0-90 degrees.

In order to determine whether the particular location on the scope face of the target-cursor intersection might bear critically on judgmental accuracy, each test item was constructed so as to appear in each of the four scope quadrants. Quadrants were designated as shown in Figure 2.

Since a principal goal of the study was to evaluate angular judgments within the framework of the NEL category system, axis angles were selected to test the various aspects of this system. Table 1 presents the upper and lower limits of the NEL categories and the test stimuli used for judgment within each.

TABLE 1

NEL Axis Angle Categories and Test Stimuli Within Each Category

Category	Test Stimuli in 1° Steps
1. 0° - 19°	9.5° - 18.5°
2. 20° - 59°	20.5° - 29.5° & 49.5° - 58.5°
3. 60° - 74°	60.5° - 73.5°
4. 75° - 84°	75.5° - 83.5°
5. 85° - 90°	85.5° - 89.5°

As can be seen from this table, there were 58 test angles. Each was presented once in each quadrant so that there was a total of 232 test items. Within the constraint that all stimuli appear in all quadrants once, the items were randomly ordered for test presentation.

Presentation of test items was by means of a Revere 35 mm, self-advancing, timed projector. Each item appeared for 12 seconds, which pre-testing indication was adequate for such judgments. Slides were numbered to assure that subjects would enter their responses on the answer sheets appropriately. The answer sheets consisted of five columns each labled with the upper and lower limits of one of the NEL categories and another column designating item number.

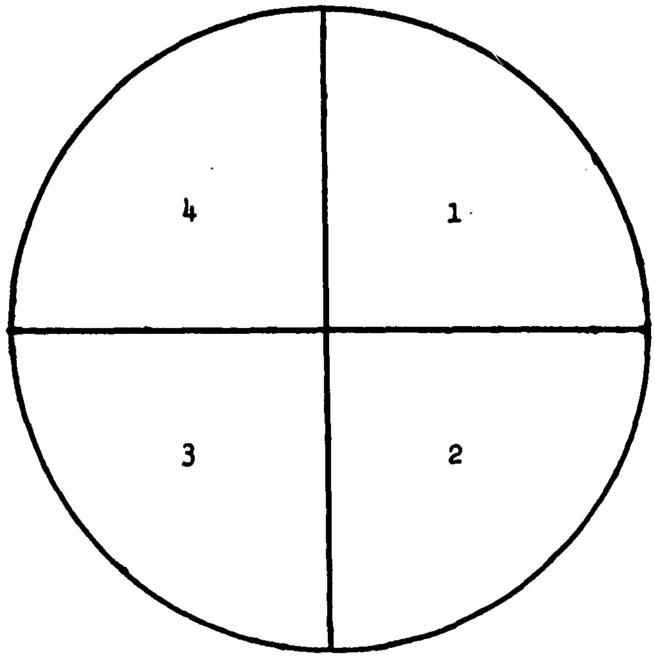

Fig. 2. Scope quadrant designations.

3. Training and Testing Procedure

The experimental groups were distinguished on the basis of whether or not they received a 30 minute training period designed to familiarize them with angular judgments. Training consisted of exposing test stimuli in 5 degree intervals from 0 to 90 degrees and from 90 to 0 degrees, and following response, providing immediate knowledge of results.

Before testing both groups received the following instructions on the task:

"Please turn to the front of your answer booklet and write your name, last name first, your rate, your class number and the date.

"You are about to see some slides that look like this.

(Project first sample slide on the screen).

"Now look at your answer sheet. You will notice that it has five columns with headings of 0 to 19 degrees; 20 to 59 degrees; 60 to 74 degrees; 75 to 84 degrees; or 85 to 90 degrees. You are to determine whether the acute angle formed by the intersection of the two straight lines, here (point to acute angle on screen) falls within the range of 0 to 19; 20 to 59; 60 to 74; 75 to 84; or 85 to 90 degrees. You will recall that an acute angle is less than 90 degrees--or less than a right angle.

"The angle you are looking at on the screen now is a 60° angle, therefore, it is in the 60 to 74 degree range. The next slide (show next sample slide) has a 30° angle. Which column should be checked for it? (Be sure that everyone understands how to mark answers.)

"The next slide you see, and those that follow it, will have a number in the upper left, here. (Point to upper left of slide on the screen.) Commencing with the next slide which is number one, please mark your answer on the answer sheet. The number of the line you check on the answer sheet must correspond with the number appearing on the screen.

"Now look at the column headings again. They are 0 to 19; 20 to 59; 60 to 74; 75 to 84; and 85 to 90 degrees. You may look at them as often as you need.

"Are there any questions?" (Commence showing slides.)

The testing of each group was carried out separately, and the total test time was divided into two equal sessions with a "smoke break" between.

C. Results

The data were initially reduced to three measures: (1) Average error of each subject in categorizing stimuli near the boundaries separating the categories, (2) Constant error in categorizing the same boundary stimuli, and (3) Percentage of judgments placing each stimulus in each category. Only stimuli near the boundaries of the system were used for measures (1) and (2) because the number of correct judgments at distances greater than 5° from the boundary were quite high and would, therefore, have added little to the analyses of the main effects.

The procedure for determining average error involved averaging responses to stimuli 1, 3, and 5 degrees from each side of each boundary for each category by means of a weighting which led to scores in arbitrary units. (See Appendix A for a description of the means for arriving at these units.) Scores ranged from a perfect score of zero, when all stimuli were correctly judged to 30, when all were incorrectly categorized. The number of category steps by which the judgments were in error was disregarded.

The procedure for determining constant error employed the same stimuli, but regrouped so as to have principal reference to the category boundaries, rather than the categories themselves. The purpose of this scoring was to obtain a measure of any predominant direction of misjudgment as a function of the principal experimental variables. The scoring procedure resulted in scores in arbitrary units ranging from 9 for crossing a boundary consistently in the direction of underestimation to 15 for boundary crossing consistently in the direction of overestimation. A score of 12 represented unbiased judgments on the average. (See Appendix B for a description of the means for arriving at these units.) The number of category steps by which the judgments were in error was disregarded.

In analyzing the percentage of subjects placing each stimulus in each category, all stimuli were employed with the exception of the one at 17.5 degrees for which there was evidence of inaccuracy in the stimulus drawing.

Table 2 presents the mean average error for each of the experimental variables, and Table 3 presents an analysis of variance of the average error data.

From the standpoint of the axis angle classification problem, the potentially important results of this analysis are the significant F values for the category, and quadrant x category interaction and the failure of significance of the group and quadrant variables.

An examination of the mean average error scores involved in the category factor (Table 2) indicates that the effect arises from the considerably greater accuracy with which the first and fifth categories are employed.

TABLE 2

Mean Average Error in Categorizing Stimuli
(Units Arbitrary, Showing Relative Position Only-See Appendix A)

Variable	Mean Average Error
Category	
1. 0° - 19°	6.54
2. 20° - 59°	13.46
3. 60° - 74°	14.44
4. 75° - 84°	13.65
5. 85° - 90°	3.67
Quadrant	
1. Upper Right	10.25
2. Lower Right	10.45
3. Lower Left	10.03
4. Upper Left	10.70
Group	
1. Trained	9.83
2. Untrained	10.85

An examination of the category means by quadrant revealed an unsystematic variation, but suggested no ready interpretation. The failure of the training and quadrant variables to reach significance has several implications which will be considered later.

TABLE 3

Analysis of Variance of the Average Error Scores

Source of Variance	Error Term	d.f.	Mean Square	F
a. Mean		1		
b. Group - Trained vs. Untrained	c	1	314.31	3.196
c. Subjects	1	58	98.34	3.282*
d. Quadrant	f	3	24.53	---
e. Quadrant x Group	f	3	15.01	---
f. Quadrant x Subjects	1	174	24.57	---
g. Category	i	4	5,793.58	4.507*
h. Category x Group	i	4	262.04	2.038
i. Category x Subjects	1	232	128.56	4.292*
j. Quadrant x Category	1	12	140.01	4.675*
k. Quadrant x Category x Group	1	12	38.07	1.271
l. Quadrant x Category x Subjects		696	29.95	---
Total		1200		

*Significant beyond the 1% level.

Table 4 presents the mean constant error for each of the experimental variables, and Table 5 presents the results of an analysis of variance of the constant error data.

TABLE 4

Mean Constant Error in Categorizing Stimuli
(Units Arbitrary-Showing Relative Position Only-See Appendix B)

Variable	Mean Constant Error
Boundary	
1. 19.5	-.08
2. 59.5	1.39
3. 74.5	1.30
4. 84.5	1.08
Quadrant	
1. Upper Right	.94
2. Lower Right	.80
3. Lower Left	1.02
4. Upper Left	.95
Group	
1. Trained	.76
2. Untrained	1.08

It is apparent that much the same set of relationships emerge from the constant error as from the average error analysis; i.e., training and quadrant fail of significance, while boundary and its interaction with quadrant are significant. From Table 4, the significant boundary effect can be seen to arise from the tendency to overestimate stimuli at boundaries 2, 3, and 4.

An examination of the boundary means by quadrant revealed an unsystematic variation which suggested no ready interpretation. The failure of significance of the training and quadrant variables will be dealt with in the next section.

TABLE 5

Analysis of Variance of the Constant Error Scores

Source of Variance	Error Term	d.f.	Mean Square	F
a. Mean		1		
b. Group-Trained vs Untrained	e	1	24.22	2.925
c. Subjects	l	58	8.28	15.623*
d. Quadrant	f	3	2.12	2.019
e. Quadrant x Group	f	3	1.76	1.676
f. Quadrant x Subjects	l	174	1.05	1.981*
g. Boundary	i	3	110.62	23.046*
h. Boundary x Group	i	3	2.32	---
i. Boundary x Subjects	l	174	4.80	9.057*
j. Quadrant x Boundary	l	9	3.57	6.736
k. Quadrant x Boundary x Group	l	9	.41	---
l. Quadrant x Boundary x Subjects		522	.53	---
Total		960		

*Significant beyond the 1% level.

Table 6 presents the mean percentage of the stimuli correctly judged over-all, and by category, for all of the test stimuli combined and for the boundary stimuli employed in the error analyses.

TABLE 6

Mean Percentage Correct Judgments
By Category and Over-all

	Category					Over-all
	1	2	3	4	5	
All Stimuli	86.3	61.4	50.2	53.1	89.2	68.0
Boundary Stimuli	72.5	49.0	49.0	51.7	86.0	61.6

It is clear that by both indices the percentage data are consistent with the average error analysis in indicating the greater accuracy of category 1 and 5 stimulus judgments over those for the three central categories.

Figure 3 presents the percentage of correct judgments for all axis angles tested (excluding 17.5 degrees which was improperly drawn). The overestimation tendency is clear in the decreasing number of correct judgments as the upper boundaries of categories 2, 3, and 4 are approached, as compared with the accuracy of the judgments at the lower boundaries of the same categories. This characteristic is not found at the lower boundary of category 2, which is consistent with the near zero constant error at this particular point.

Neither the average nor constant error analyses considered the extent of miscategorization error. To get at this, the percentage data were tabulated to indicate the distribution of error frequencies by extent of the error. The analysis clearly indicated that one category displacements constituted far and away the dominant source of all errors--making up 92% of them, while two category errors were found in only 7% of the judgments, and larger errors made up less than 1/100th of the total.

D. Discussion

The failure of the 30 minute familiarization training period to produce reliable differences in the accuracy of axis angle judgments is of importance in any system such as NEL's because, it may either mean that the population expected to make these judgments comes to the problem with an informally acquired ability to perform without formalized instruction, or that the very brief

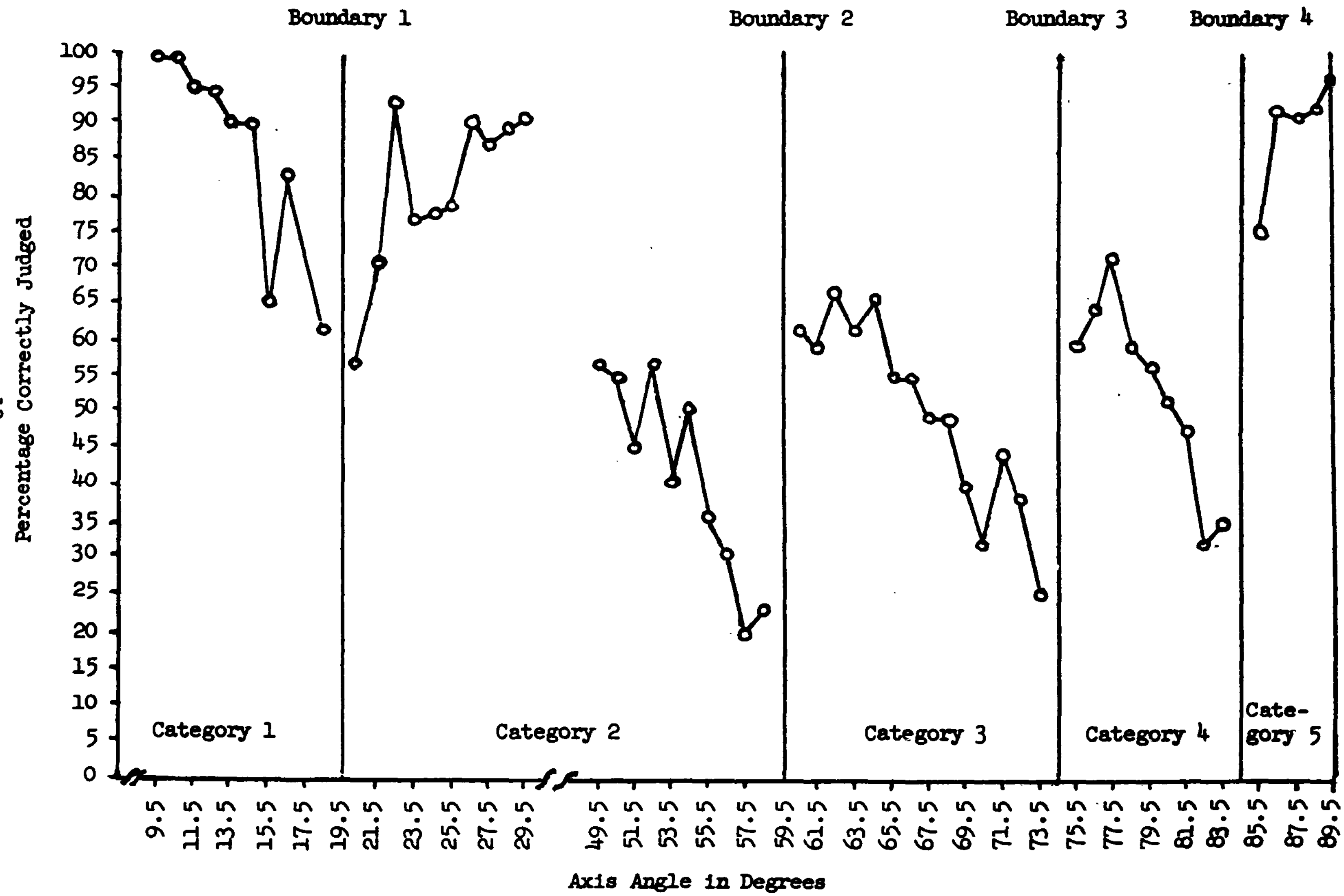

Fig. 3. Percentage Correct Judgments by Axis Angle

duration and nature of the training were inadequate. Though in both the average and constant error analyses the training factor was insignificant, the direction of the difference was in the expected direction. It is quite possible that a more extensive training program, aimed at curtailing the pattern of constant errors found in these data, could well have a greater influence; particularly when the added load of determining the target's major axis from the more ambiguous stimuli of a real PPI scope is involved.

The failure of the scope quadrant variable even to approach significance is certainly a welcome result from the operational standpoint, since it means that neither over-all accuracy nor predominant error direction should be expected to be influenced by the particular scope location in which a target may appear. There is no reason to expect this result not to be applicable to real PPI scope material.

The constant error analysis is more nearly in accord with those investigators who report a predominant tendency twoard angular overestimation than those reporting the reverse; the exception being near the most acute angular category where there is virtually no constant error. As suggested above, this result provides direction in preparing training materials pertaining to this type of judgment. As with quadrant effects, there is no reason to suppose that these constant error relationships are not applicable to the operational stimulus situation.

The significant boundary and category effects are difficult to interpret because of a couple of artifacts inherent in the category system used. First there is the range constraining effect of restricting judgments to $\frac{1}{4}$ of a complete circle of angles. As a result, it was not possible either to obtain a negative error on category 1 stimuli or a positive error on category 5 stimuli. To some extent, this accounts for the lower error rates within those categories, as well as the pattern of constant error scores at their boundaries. A second artifact is the variable size of the categories, which likely produced some undetermined distortion on all measures. This is not to imply that the operational usefulness of the conclusions regarding category differences is thereby reduced. On the contrary these data reflect the effects to be expected within the proposed axis angle category system, and are directly applicable to any developments or modifications in the system which might be contemplated as a result.

Without an extensive study of the influence of various levels of average and constant errors at various dividing points in the NEL categories, it is difficult to assess the significance of the percentage correct judgment data on final target classification. Whether a low point of between 50% and 60% correct categorizing

in the central categories is too great for the system's target classification logic to tolerate cannot at this time be stated--particularly since the extent of degradation beyond this with PPI scope materials is still unknown. If this accuracy is inadequate, then either some angle estimation assist device may have to be developed, or a more extensive training program, aimed at the sorts of errors revealed in this research, will be needed.

E. Summary and Conclusions

In order to assess the accuracy with which observers can be expected to make certain angular judgments of importance in target classification by scanning sonar, a study was carried out utilizing line-drawing, idealized simulations of PPI scope presentations of target axis angles. The subjects were required to estimate the angles by means of the NEL category system. One group of subjects was given 30 minutes of familiarization training, while the other group was simply introduced to the task.

The following conclusions can be drawn:

1. Both in terms of average and constant error, neither the training received nor the quadrant of the scope in which the test angles appeared had a significant influence on accuracy.

2. Both in terms of average and constant error, accuracy was influenced significantly by the angular category of the test stimuli; being higher at the extreme upper and lower categories. Two artifacts in the category system which may well have influenced this result were pointed out.

3. The dominant tendency was for axis angles to be over-estimated, except at the most acute angular category, where little constant error was found.

4. The percentage of correct judgments ranged from 50.2% to 89.2%, depending upon category. The over-all accuracy was 68.0%. The great majority of misclassifications (92.0%) were of only one category displacement in extent.

5. The failure of the training here used to produce significant effects indicates an approach to training in terms of dominant constant error tendencies.

REFERENCES

1. Barden, H. Trekk av synsopfatningen og dens utvekling. Et bidrag til overgangspolderens psykologi. (Features of visual perception and its development: A contribution to the psychology of adolescence). Oslo: Jacob Dybwatt, 1933 (Psych. Abstr. 8:3387).

2. Calabresi, R. Osservazioni sulla valutazione tattilo-cinetica degli angoli. (On the tactual-kinetic estimation of angles). Arch ital. di. psicol., 1933, 11, (Psych. Abstr. 8:45).

3. Gales, R.S., & Eady, H.R. The human factor in sonar detection and classification. U.S. Navy Journal of Underwater Acostics, 1958, 8. (CONFIDENTIAL).

4. Mackie, R.R., & Kimmel, H.D. Research on sonar target classification using the AN/SQS-10 PPI scope. Human Factors Research, Inc., Los Angeles, California, 1954. (For Personnel and Training Branch Office of Naval Research, Dept. of the Navy), Contract Nonr 2649(00). (CONFIDENTIAL).

5. Pratt, M.B. The visual estimation of angles. J. exp. Psychol., 1926, 9.

6. Reese, E.P., Reese, T.W., Volkmann, J., & Corbin, H.H. Psychophysical research; summary report 1946-1952. U.S.N. Spec. Dev. Cent. Tech. Rep., 1953, S.D.C.-131-1-5.

7. Rogers, S. The anchoring of absolute judgments. Arch. Psychol., N.Y., 1941, No. 261.

8. Saltzman, D.C., & Garner, W.R. Accuracy of visual estimation of azimuth position. J. Psychol., 1950, 29.

9. Samanta, M.N. Visual estimation of angles. Indian J. Psychol., 1928, 3. (Psych. Abstr. 3:2956).

10. Stavrianos, B.K. The relation of shape preception to explicit judgments of inclination. Arch. Psychol., N.Y., 1945, No. 296.

11. Takala, Mortti. On constant errors in the judgment of the degree of inclination. Acta Psychol. Fennica, 1951, 1. (Psych. Abstr. 29:6642).

APPENDIX A

PROCEDURES FOR COMPUTING MEAN AVERAGE ERROR

Mean average error per category was obtained by utilizing stimuli 1, 3, and 5 degrees from the upper and lower category boundaries. Since category 1 has only an upper boundary within the 90 degree arc, and category 5 only a lower boundary, the results from only 3 stimuli each were employed in their averages. On the other hand, categories 2 and 3 both have two boundaries; i.e., upper and lower, so that judgments of 6 stimuli each were employed in their averages. Although category 4 has both an upper and a lower boundary, the small size of the category restricted the number of stimuli to be averaged to 5. Because of the differing number of stimulus judgments entering into the mean average error; i.e., 3 for categories 1 and 5; 6 for categories 2 and 3; and 5 for category 4, the common denominator of 3, 5, and 6; i.e., 30 was employed to get a comparable value.

Specifically, if an observer got all test stimuli correct in any category he received a score of 30. If he miscategorized all stimuli, he scored zero. Each error in categories 1 and 5 subtracted 10 points from the total possible of 30. Each error in categories 2 and 3 subtracted 5 points, and each error in category 4 subtracted 6 points.

APPENDIX B

PROCEDURES FOR COMPUTING MEAN CONSTANT ERROR

Direction of error tendency was of principal interest in this analysis so that the arbitrary unit scoring system emphasized this aspect. Stimuli 1, 3, and 5 degrees above and below each of the boundaries; i.e., 19.5, 58.5, 74.5 and 84.5 degrees, were utilized. This resulted in 6 stimuli being employed in the total score at each boundary. A score of 2 was assigned all judgments showing no constant error, a score of 1 to those showing underclassification, and a score of 3 to those showing overclassification. Thus, if a stimulus near the upper boundary of a category was judged as falling in that category or lower, a score of 2 was recorded. Likewise, if a stimulus near the lower boundary of a category was judged as falling in that category or higher, a score of 2 was recorded. However, any stimulus near an upper boundary which was judged as falling in a higher category was scored 3, while any stimulus near a lower boundary which was judged as falling in a lower category was scored 1.

Since there are only 5 stimuli available in category 4 in 2 degree steps, the stimulus in the middle of this category; i.e., 79.5 degrees, was scored twice for each subject; once as a lower and once as an upper boundary stimulus.